

# BULK BEHAVIOUR OF SCHUR-HADAMARD PRODUCTS OF SYMMETRIC RANDOM MATRICES

ARUP BOSE \*

Statistics and Mathematics Unit

Indian Statistical Institute

203 B. T. Road, Kolkata 700108

INDIA

E-mail: bosearu@gmail.com

SOUMENDU SUNDAR MUKHERJEE

Master of Statistics student

Indian Statistical Institute

203 B. T. Road, Kolkata 700108

INDIA

E-mail: soumendu041@gmail.com

First version February 10, 2014

This version March 12, 2014

## Abstract

We develop a general method for establishing the existence of the Limiting Spectral Distributions (LSD) of Schur-Hadamard products of independent symmetric patterned random matrices. We apply this method to show that the LSDs of Schur-Hadamard products of some common patterned matrices exist and identify the limits. In particular, the Schur-Hadamard product of independent Toeplitz and Hankel matrices has the semi-circular LSD. We also prove an invariance theorem that may be used to find the LSD in many examples.

**Key words and phrases.** Patterned matrices, Schur-Hadamard product, limiting spectral distribution, Toeplitz, Wigner, Hankel, Circulant matrices, semi-circular law.

**AMS 2010 Subject Classifications.** Primary 15B52, 60B20; secondary 60B10, 60F99, 60B99.

---

\*Research supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India.

# 1 Introduction

Let  $A_n$  be an  $n \times n$  matrix with eigenvalues  $\lambda_1, \dots, \lambda_n$ . The empirical spectral measure  $\mu_n$  of  $A_n$  is the random measure

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}, \quad (1.1)$$

where  $\delta_x$  is the Dirac delta measure at  $x$ . The corresponding random probability distribution function is known as the *Empirical Spectral Distribution* (ESD) and denoted by  $F^{A_n}$ . The sequence  $\{F^{A_n}\}$  is said to converge (weakly) almost surely to a non-random distribution function  $F$  if, outside a null set, as  $n \rightarrow \infty$ ,  $F^{A_n}(\cdot) \rightarrow F(\cdot)$  at all continuity points of  $F$ .  $F$  is known as the *Limiting Spectral Distribution* (LSD). If the latter convergence is in probability, then the weak convergence is said to hold in probability.

There has been a lot of recent work on obtaining the LSDs of large dimensional patterned random matrices. These matrices may be defined as follows. Let  $\{x_i, x_{i,j} \mid i, j \geq 0\}$  be a sequence of random variables, called an *input sequence*. Let  $\mathbb{Z}$  be the set of all integers and let  $\mathbb{Z}_+$  be the set of all non-negative integers. Let

$$L_n : \{1, 2, \dots, n\}^2 \rightarrow \mathbb{Z}_+ \text{ (or } \mathbb{Z}_+^2) \text{ } n \geq 1, \quad (1.2)$$

be a sequence of functions. We write  $L_n = L$  and call it the *link* function and by abuse of notation we write  $\mathbb{Z}_+^2$  as the common domain of  $\{L_n\}$ . Matrices of the form

$$A_n = n^{-1/2}((x_{L(i,j)})) \quad (1.3)$$

are called *patterned random matrices*. If  $L(i, j) = L(j, i)$  for all  $i, j$ , then the matrix is symmetric. In this article we shall denote the LSD of  $\{n^{-1/2}A_n\}$ , if it exists, by  $\mathcal{L}_A$ .

There are a host of LSD results for real symmetric patterned random matrices. See, for example, [Bose and Sen \[2008\]](#) for a detailed description. The symmetric patterned matrices that have received particular attention in the literature are the Wigner, Toeplitz, Hankel, Reverse Circulant and the Symmetric Circulant matrices. Their link functions is given in Table 1.

LSD existence is also known for the upper triangular versions of these matrices [\[Basu et al., 2012\]](#). Joint convergence in terms of convergence of moments of all polynomials of these matrices has also been established in varying degrees (see, for example, [Bose et al. \[2011\]](#); [Basu et al. \[2012\]](#)).

However, the Schur-Hadamard (entrywise) product of such matrices does not seem to have been dealt with in any systematic manner. Such matrices have come up in the random matrix literature in specific situations. For example, [Bai and Zhang \[2007\]](#) considered the problem of finding the LSD of a sparse sample covariance  $S$ -matrix. They modeled the sparsity by taking Schur-Hadamard product with a

Matrix	Notation	Link function
Wigner	$W_n$	$L_W(i, j) = (\min\{i, j\}, \max\{i, j\})$
Toeplitz	$T_n$	$L_T(i, j) =  i - j $
Hankel	$H_n$	$L_H(i, j) = i + j$
Symmetric Circulant	$SC_n$	$L_{SC}(i, j) = \frac{n}{2} -  \frac{n}{2} -  i - j  $
Reverse Circulant	$RC_n$	$L_{RC}(i, j) = (i + j)(\bmod n)$
Doubly Symmetric Hankel	$DH_n$	$L_{DH}(i, j) = \frac{n}{2} -  \frac{n}{2} - (i + j)(\bmod n) $

Table 1: Some common symmetric patterned matrices and their link functions.

sparse 0-1 Wigner matrix and established the semi-circular law as the LSD of the resulting sparse  $S$ -matrix under Lindeberg-type conditions on the matrix-entries.

More recently, [Beckwith et al. \[2011\]](#) considered Schur-Hadamard product of a  $\pm 1$  Bernoulli( $p$ ) Wigner matrix with a Toeplitz matrix and established the existence of the LSD. They found that when  $p = 0.5$ , the LSD is the familiar semi-circular law, but when  $p \neq 0.5$ , the limiting moments are polynomials in  $(2p - 1)$  whose coefficients could not be identified, and as  $p$  approaches 1 these moments approach the corresponding moments of the LSD of the Toeplitz matrix. [Goldmakher et al. \[2013\]](#) considered randomly weighted sequences of  $d$ -regular graphs with size growing to  $\infty$ , which amounts to taking Schur-Hadamard product of random real symmetric weight matrices with the adjacency matrices of the graphs, and established the existence of a limiting spectral distribution that depends only on  $d$  and the distribution of the weights, under the usual decay condition on the number of  $k$ -cycles relative to the graph-size, for each  $k \geq 3$  (in the unweighted case, the limiting spectral distribution is the well-known Kesten's measure).

In this article we shall consider Schur-Hadamard products of real symmetric patterned matrices and establish results on their LSD. In particular, we prove an invariance theorem which yields the result of [Beckwith et al. \[2011\]](#), when  $p = 0.5$ , as a special case. We also consider the Schur-Hadamard product of Toeplitz and Hankel matrices (and other combinations like Toeplitz and Reverse Circulant etc.) and show that the LSD is the semi-circular law. Table 2 summarizes our results about the six patterned matrices mentioned in Table 1.

$X_n$	$Y_n$	LSD
$W_n$	$T_n, H_n, SC_n, RC_n, DH_n$	$\mathcal{L}_W$
$T_n, SC_n$	$H_n, RC_n, DH_n$	$\mathcal{L}_W$
$T_n$	$SC_n$	$\mathcal{L}_T$
$H_n$	$RC_n, DH_n$	$\mathcal{L}_H$
$RC_n$	$DH_n$	$\mathcal{L}_{RC}$

Table 2: LSDs of several Schur-Hadamard products.

## 2 Preliminaries

We shall use the method of moments to establish the existence of the LSD. For any matrix  $A$ , let  $\beta_h(A)$  denote the  $h$ -th moment of the ESD of  $A$ . The following lemma, which is easy to prove, will be useful.

**Lemma 2.1.** *Let  $\{A_n\}$  be a sequence of random matrices with all real eigenvalues. Suppose there exists a sequence  $\beta_h$  such that*

(i) *For every  $h \geq 1$ ,  $\mathbb{E}(\beta_h(A_n)) \rightarrow \beta_h$ ,*

(ii)  *$\text{Var}(\beta_h(A_n)) \rightarrow 0$  for every  $h \geq 1$  and*

(iii) *the sequence  $\{\beta_h\}$  satisfies Carleman's condition,  $\sum \beta_{2h}^{-1/2h} = \infty$ .*

*Then the LSD of  $F^{A_n}$  exists in probability and equals  $F$  with moments  $\{\beta_h\}$ . If in place of (ii),  $\beta_h$  satisfies the stronger condition*

(ii')  *$\sum_{n=1}^{\infty} \mathbb{E}[\beta_h(A_n) - \mathbb{E}(\beta_h(A_n))]^4 < \infty$  for every  $h \geq 1$ ,*

*then the LSD exists in the almost sure sense.*

We shall consider three assumptions on the input sequence.

**(A1).** The input random variables are independent and uniformly bounded with mean 0, and variance 1.

**(A2).** The input random variables are i.i.d. with mean 0 and variance 1.

**(A3).** The input random variables are independent with mean 0 and variance 1, and with uniformly bounded moments of all orders.

We now quickly recollect some terminology and notation from the general theory of patterned matrices (see [Bose and Sen \[2008\]](#)).

A link function  $L$  is said to satisfy *Property B* if

$$\Delta_L := \sup_n \sup_{t \in \text{range}(L)} \sup_{1 \leq k \leq n} \#\{l \mid 1 \leq l \leq n, L(K, l) = t\} < \infty.$$

In other words, the total number of times any particular variable appears in any row is uniformly bounded. All the matrices introduced so far satisfy this property. For example,  $\Delta_{L_W} = 1$  and  $\Delta_{L_T} = 2$ .

Let  $L$  be the link function of the matrix  $A_n$ . Define

$$k_n^A := \#\{L_n(i, j) \mid 1 \leq i, j \leq n\},$$

and

$$\alpha_n^A := \max_k \#\{(i, j) \mid L_n(i, j) = k\}.$$

Consider the following conditions on  $L$ :

$$k_n^A \rightarrow \infty \quad \text{and} \quad k_n^A \alpha_n^A = O(n^2), \quad (2.1)$$

where  $O(\cdot)$  is the Landau big “Oh” notation: for two real valued functions  $f$  and  $g$  defined on the set of integers, one writes  $f(n) = O(g(n))$  if there is a constant  $C$  independent of  $n$  such that  $|f(n)| \leq C|g(n)|$  for all  $n > N_0$  for some integer  $N_0$ . Often the constant  $C$  and the “cut-off”  $N_0$  depend on other “parameters”, say  $\alpha$ , of the problem at hand, and one makes this explicit by writing  $f(n) = O_\alpha(g(n))$ . For instance, we might have written in (2.1) that  $k_n^A \alpha_n^A = O_L(n^2)$ , because the constant implied by the big “Oh” might depend on the link function. However, from now on we shall suppress these “parameters” to avoid notational clutter. Note that all the link functions introduced so far satisfy the above conditions. It is known that if the LSD of  $\{n^{-1/2}A_n\}$  exists under Assumption (A1), then the same LSD continues to hold under (A2) or (A3), provided the link function satisfies Property B and Conditions (2.1). The same continues to be true for Schur-Hadamard products. Thus, in our arguments, without loss of any generality, we assume that (A1) holds. Traditionally, LSD results are stated under (A1), and (A3) is appropriate while studying the joint convergence of more than one sequence of matrices.

The *Moment-Trace Formula* plays a key role in this approach. A function

$$\pi : \{0, 1, \dots, h\} \rightarrow \{1, 2, \dots, n\}$$

with  $\pi(0) = \pi(h)$  is called a *circuit* of length  $h$ . The dependence of a circuit on  $h$  and  $n$  is suppressed. Then

$$\beta_h(A) = \frac{1}{n} \text{tr}(A^h) = \frac{1}{n} \sum_{\pi \text{ circuit of length } h} a_\pi, \quad (2.2)$$

where

$$a_\pi := a_{L(\pi(0), \pi(1))} a_{L(\pi(1), \pi(2))} \cdots a_{L(\pi(h-1), \pi(h))}.$$

If  $L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))$ , with  $i < j$ , we shall use the notation  $(i, j)$  to denote such a match of the  $L$ -values. Also if  $L$  is the link function of  $A_n$ , we will often use the further shorthand notation  $i \sim^A j$  in lieu of the phrase “ $(i, j)$  is an  $L$ -match”. If an  $L$ -value is repeated exactly  $e$  times, we say that the circuit  $\pi$  has an edge (or  $L$ -edge) of order  $e$  ( $1 \leq e \leq h$ ). If  $\pi$  has all  $e \geq 2$ , then it is called  $L$ -matched (in short matched). If  $\pi$  has only order two edges, then it is called pair-matched. From the general theory, it follows that only pair-matched circuits are relevant when computing limits of moments.

To deal with Conditions (ii) and (ii') of Lemma 2.1, we need multiple circuits:  $t$  circuits  $\pi_1, \pi_2, \dots, \pi_t$  are *jointly  $L$ -matched* if each  $L$ -value occurs at least twice across all circuits. They are *across  $L$ -matched* if each circuit has at least one  $L$ -value which occurs in at least one of the other circuits.

Two circuits  $\pi_1$  and  $\pi_2$  are equivalent if and only if their  $L$ -values respectively match at the same locations, i.e., if for all  $i, j$ ,

$$L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Leftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi_2(j)).$$

Any equivalence class can be indexed by a partition of  $\{1, 2, \dots, h\}$ . We label these partitions by *words* of length  $h$  of letters where the first occurrence of each letter is in alphabetical order. For example, if  $h = 4$  then the partition  $\{\{1, 3\}, \{2, 4\}\}$  is represented by the word *abab*. This identifies all circuits  $\pi$  for which  $L(\pi(0), \pi(1)) = L(\pi(2), \pi(3))$  and  $L(\pi(1), \pi(2)) = L(\pi(3), \pi(1))$ . Let  $w[i]$  denote the  $i$ -th entry of  $w$ . The equivalence class corresponding to  $w$  is

$$\Pi(w) := \{\pi \mid w[i] = w[j] \Leftrightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\}.$$

Note that the number of partition blocks corresponding to  $w$  is same as the number of distinct letters in  $w$ , which we denote by  $|w|$ . By varying  $w$ , we obtain all the equivalence classes. It is important to note that for any fixed  $h$ , even as  $n \rightarrow \infty$ , the number of words remains finite but the number of circuits in any given  $\Pi(w)$  may grow indefinitely. Henceforth, we shall denote the set of all words of length  $h$  by  $\mathcal{A}_h$ . Notion of matches carry over to words and we shall again use the notation  $(i, j)$ ,  $i < j$  to denote the match  $w[i] = w[j]$  in  $w$ . Note that a word is *pair-matched* if every letter appears exactly twice in that word. The set of all pair-matched words of length  $2k$  is denoted by  $\mathcal{W}_{2k}$ . For technical reasons it is often easier to deal with a class larger than  $\Pi(w)$ :

$$\Pi^*(w) = \{\pi \mid w[i] = w[j] \Rightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\}.$$

Any  $i$  (or  $\pi(i)$  by abuse of notation) is a *vertex*. It is *generating* if either  $i = 0$  or  $w[i]$  is the first occurrence of a letter. Otherwise, it is called non-generating. For example, if  $w = abbcab$  then  $\pi(0), \pi(1), \pi(2), \pi(4)$  are generating and  $\pi(3), \pi(5), \pi(6)$  are non-generating. By Property B a circuit is completely determined, up to finitely many choices, by its generating vertices. The number of generating vertices in any circuit in  $\Pi(w)$  is  $|w| + 1$  and hence

$$\#\Pi(w) \leq \#\Pi^*(w) = O(n^{|w|+1}).$$

The set of generating vertices (indices) is denoted by  $S$ . The dependence on the word  $w$  will, in general, be clear from the context. Sometimes we shall write  $\Pi_A(w)$ ,  $\Pi_A^*(w)$  or  $S_A$  to emphasise dependence on the matrix  $A_n$ .

From the general theory, it follows that for a sequence of patterned random matrices  $\{n^{-1/2}A_n\}$  the LSD exists if for all  $w \in \mathcal{W}_{2k}$ , the following limit exists:

$$p(w) = \lim n^{-(1+k)} \#\Pi(w) = \lim n^{-(1+k)} \#\Pi^*(w) \quad (2.3)$$

and in that case the  $2k$ -th moment of the LSD is given by

$$\beta_{2k} = \sum_{w \in \mathcal{W}_{2k}} p(w).$$

The existence of the LSD for the Wigner, Hankel, Toeplitz, Reverse Circulant, Symmetric Circulant and Doubly Symmetric Hankel matrices can be established by verifying that for every  $k$ ,  $p(w)$  exists for every pair-matched word (see [Bose and Sen \[2008\]](#); [Bose et al. \[2010\]](#)).

In the next section we extend the above approach for a single sequence to the Schur-Hadamard product of two sequences. It may be noted that although we work with only two sequences, it is quite straightforward to extend the results to any finite number of sequences.

### 3 Schur-Hadamard product

Suppose  $\{X_n\}$  and  $\{Y_n\}$  are two independent sequences of patterned symmetric random matrices with link functions  $L_X$  and  $L_Y$  respectively and their input sequences satisfy (A1). Let  $Z_n = X_n \odot Y_n$  be their Schur-Hadamard product. We shall employ the moment method via the word approach to study the LSD of  $\{n^{-1/2}Z_n\}$ . Note that  $Z_n$  is not necessarily a patterned matrix. However, many of the arguments of the general theory for a single matrix may be used for  $Z_n$  with appropriate modifications.

Define

$$k_n^Z := \#\{(L_X(i, j), L_Y(i, j)) \mid 1 \leq i, j \leq n\},$$

i.e.,  $k_n^Z$  is the total number of the  $X$  and  $Y$  variable pairs appearing in the matrix  $Z_n$ . Also let

$$\alpha_n^Z := \max_{(k, l)} \#\{(i, j) \mid (L_X(i, j), L_Y(i, j)) = (k, l)\},$$

i.e.,  $\alpha_n^Z$  is the maximum number of occurrences of any  $X$  and  $Y$  variable pair in  $Z_n$ .

It is easy to show that if  $L_X$  and  $L_Y$  both satisfy (2.1), then  $\{k_n^Z, \alpha_n^Z\}$  also satisfy these conditions.

To elaborate, note that

$$\max\{k_n^X, k_n^Y\} \leq k_n^Z \leq k_n^X + k_n^Y.$$

So  $k_n^Z \rightarrow \infty$ .

Similarly, it is obvious that

$$\alpha_n^Z \leq \min\{\alpha_n^X, \alpha_n^Y\}.$$

Therefore

$$\begin{aligned} k_n^Z \alpha_n^Z &\leq (k_n^X + k_n^Y) \min\{\alpha_n^X, \alpha_n^Y\} \\ &\leq \alpha_n^X k_n^X + \alpha_n^Y k_n^Y \\ &= O(n^2) + O(n^2) = O(n^2), \end{aligned}$$

which completes the verification of (2.1).

For the time being assume *all moments exist*. Then the *moment trace formula* for  $n^{-1/2}Z_n$  becomes

$$\begin{aligned}\beta_h(F^{n^{-1/2}Z_n}) &= n^{-(1+\frac{h}{2})} \sum_{\pi: \pi \text{ circuit of length } h} z_\pi \\ &= n^{-(1+\frac{h}{2})} \sum_{\pi: \pi \text{ circuit of length } h} x_\pi y_\pi.\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{E}(\beta_h(F^{n^{-1/2}Z_n})) &= n^{-(1+\frac{h}{2})} \sum_{\pi: \pi \text{ circuit of length } h} \mathbb{E}(x_\pi y_\pi) \\ &= n^{-(1+\frac{h}{2})} \sum_{\pi: \pi \text{ circuit of length } h} \mathbb{E}x_\pi \mathbb{E}y_\pi.\end{aligned}$$

The two (possibly different) link functions  $L_X$  and  $L_Y$  induce two partitions (via words)  $\mathcal{P}_X$  and  $\mathcal{P}_Y$  on the set of all circuits of length  $h$ . Consider the resultant  $\mathcal{P}_Z$  of these two partitions defined as

$$\mathcal{P}_Z = \{\Pi_X(w) \cap \Pi_Y(w') \mid w, w' \in \mathcal{A}_h\}.$$

For the sake of brevity, let us define

$$\Pi_Z(w, w') := \Pi_X(w) \cap \Pi_Y(w').$$

Then we can write

$$\mathbb{E}(\beta_h(F^{n^{-1/2}Z_n})) = n^{-(1+\frac{h}{2})} \sum_{(w, w') \in \mathcal{A}_h^2} \sum_{\pi \in \Pi_Z(w, w')} \mathbb{E}x_\pi \mathbb{E}y_\pi. \quad (3.1)$$

We can now state and prove our first lemma.

**Lemma 3.1.** *Suppose the input sequences satisfy Assumption (A1) and the link functions  $L_X$  and  $L_Y$  satisfy Property B. Then circuits which have at least one edge of order  $\geq 3$  contribute zero to the possible limit of  $\mathbb{E}(\beta_h(F^{n^{-1/2}Z_n}))$ . As a consequence, for every odd  $h$ ,  $\mathbb{E}(\beta_h(F^{n^{-1/2}Z_n})) \rightarrow 0$  and when  $h$  is even, only those words  $(w, w')$  where both are pair-matched can contribute to the possible limit of moments.*

*Proof.* First note that if a circuit  $\pi$  is not  $L_X$ -matched or  $L_Y$ -matched then  $\mathbb{E}x_\pi \mathbb{E}y_\pi = 0$  and consequently such a  $\pi$  does not have any contribution to the moment. We thus need to consider only matched circuits (or words) henceforth. Denote by  $C_{h,3+}^L$  the set of all  $L$ -matched circuits of length  $h$  with at least one edge of order  $\geq 3$ . If  $L$  satisfies Property B, then Lemma 1(a) of [Bose and Sen \[2008\]](#) ensures that

$$\#C_{h,3+}^L = O(n^{\lfloor (h+1)/2 \rfloor}), \quad (3.2)$$



where  $\lfloor x \rfloor$  denotes the largest integer contained in  $x$ . Using this in our context we have

$$\begin{aligned} \#(C_{h,3+}^{L_X} \cup C_{h,3+}^{L_Y}) &\leq \#C_{h,3+}^{L_X} + \#C_{h,3+}^{L_Y} \\ &= O(n^{\lfloor (h+1)/2 \rfloor}) + O(n^{\lfloor (h+1)/2 \rfloor}) \\ &= O(n^{\lfloor (h+1)/2 \rfloor}). \end{aligned}$$

This implies that the circuits which have at least one  $L_X$ -edge or  $L_Y$ -edge of order  $\geq 3$  do not contribute in the limit. It follows immediately that for odd  $h$  the limit of the expected  $h$ -th moment is zero and for even  $h$  only the circuits in  $\Pi_Z(w, w')$  where both  $w$  and  $w'$  are pair-matched can have a potential contribution. This proves the lemma.  $\square$

**Lemma 3.2.** *Suppose  $L_X$  and  $L_Y$  satisfy Property B and the input sequences satisfy Assumption (A1). Then  $\{\beta_h(n^{-1/2}Z_n)\}$  satisfies Condition (ii) of Lemma 2.1 for any  $h$ .*

*Proof.* Let  $K_{h,t}^L$  be the set of  $t$ -tuples of circuits  $(\pi_1, \dots, \pi_t)$  of length  $h$  such that they are jointly and across  $L$ -matched. Using Lemma A.1 we have

$$\begin{aligned} \#(K_{h,2}^{L_X} \cup K_{h,2}^{L_Y}) &\leq \#K_{h,2}^{L_X} + \#K_{h,2}^{L_Y} \\ &= O(n^{h+1}) + O(n^{h+1}) \\ &= O(n^{h+1}). \end{aligned} \tag{3.3}$$

Now write

$$\begin{aligned} \text{Var}(\beta_h(n^{-1/2}Z_n)) &= \mathbb{E}[\beta_h(n^{-1/2}Z_n) - \mathbb{E}(\beta_h(n^{-1/2}Z_n))]^2 \\ &= \mathbb{E}[n^{-1} \text{tr}(n^{-1/2}Z_n)^h - \mathbb{E}(n^{-1} \text{tr}(n^{-1/2}Z_n)^h)]^2 \\ &= \frac{1}{n^{h+2}} \mathbb{E}[\text{tr} Z_n^h - \mathbb{E} \text{tr} Z_n^h]^2 \\ &= \frac{1}{n^{h+2}} \sum_{(\pi_1, \pi_2)} \mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}), \end{aligned} \tag{3.4}$$

and decompose

$$\begin{aligned} z_{\pi_j} - \mathbb{E}z_{\pi_j} &= x_{\pi_j}y_{\pi_j} - \mathbb{E}x_{\pi_j}\mathbb{E}y_{\pi_j} \\ &= (x_{\pi_j} - \mathbb{E}x_{\pi_j})y_{\pi_j} + (y_{\pi_j} - \mathbb{E}y_{\pi_j})\mathbb{E}x_{\pi_j}. \end{aligned} \tag{3.5}$$

Note that by decomposition (3.5) we have

$$\begin{aligned} &\mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}) \\ &= \mathbb{E}((x_{\pi_1} - \mathbb{E}x_{\pi_1})y_{\pi_1} + (y_{\pi_1} - \mathbb{E}y_{\pi_1})\mathbb{E}x_{\pi_1})((x_{\pi_2} - \mathbb{E}x_{\pi_2})y_{\pi_2} + (y_{\pi_2} - \mathbb{E}y_{\pi_2})\mathbb{E}x_{\pi_2}) \\ &= \mathbb{E}(x_{\pi_1} - \mathbb{E}x_{\pi_1})(x_{\pi_2} - \mathbb{E}x_{\pi_2})\mathbb{E}y_{\pi_1}y_{\pi_2} + \mathbb{E}(x_{\pi_1} - \mathbb{E}x_{\pi_1})\mathbb{E}x_{\pi_2}\mathbb{E}y_{\pi_1}(y_{\pi_2} - \mathbb{E}y_{\pi_2}) \\ &\quad + \mathbb{E}x_{\pi_1}\mathbb{E}(x_{\pi_2} - \mathbb{E}x_{\pi_2})\mathbb{E}(y_{\pi_1} - \mathbb{E}y_{\pi_1})y_{\pi_2} + \mathbb{E}x_{\pi_1}\mathbb{E}x_{\pi_2}\mathbb{E}(y_{\pi_1} - \mathbb{E}y_{\pi_1})(y_{\pi_2} - \mathbb{E}y_{\pi_2}) \\ &= \mathbb{E}(x_{\pi_1} - \mathbb{E}x_{\pi_1})(x_{\pi_2} - \mathbb{E}x_{\pi_2})\mathbb{E}y_{\pi_1}y_{\pi_2} + \mathbb{E}x_{\pi_1}\mathbb{E}x_{\pi_2}\mathbb{E}(y_{\pi_1} - \mathbb{E}y_{\pi_1})(y_{\pi_2} - \mathbb{E}y_{\pi_2}). \end{aligned} \tag{3.6}$$

If  $(\pi_1, \pi_2)$  are not jointly  $L_X$ -matched, then one of the circuits, say  $\pi_1$ , has an  $L_X$ -value which does not occur anywhere else. Therefore  $\mathbb{E}x_{\pi_1} = 0$ . So, from (3.6) it follows that

$$\mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}) = \mathbb{E}x_{\pi_1}(x_{\pi_2} - \mathbb{E}x_{\pi_2})\mathbb{E}y_{\pi_1}y_{\pi_2}.$$

But since the input  $X$ -variable corresponding to the single  $L_X$ -value appears in the product  $x_{\pi_1}(x_{\pi_2} - \mathbb{E}x_{\pi_2})$  inside  $x_{\pi_1}$  and is independent of every other term in the product, we conclude that

$$\mathbb{E}x_{\pi_1}(x_{\pi_2} - \mathbb{E}x_{\pi_2}) = 0,$$

and as a result

$$\mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}) = 0.$$

Therefore, it is enough to consider those  $(\pi_1, \pi_2)$  which are jointly  $L_X$ -matched and jointly  $L_Y$ -matched.

Now suppose that  $(\pi_1, \pi_2)$  are jointly  $L_X$  as well as  $L_Y$ -matched but neither across  $L_X$ -matched nor across  $L_Y$ -matched. Then there is a circuit, say  $\pi_k$ , which is only *self*  $L_X$ -matched, i.e., none of its  $L_X$ -values is shared with those of the other circuit. Similarly, there is a circuit  $\pi_l$  that is only self  $L_Y$ -matched. Now note that  $(x_{\pi_k} - \mathbb{E}x_{\pi_k})$  is independent of  $(x_{\pi_j} - \mathbb{E}x_{\pi_j})$  for  $j \neq k$  and similarly  $(y_{\pi_l} - \mathbb{E}y_{\pi_l})$  is independent of  $(y_{\pi_j} - \mathbb{E}y_{\pi_j})$  for  $j \neq l$ . Using this in (3.6) we can write

$$\begin{aligned} \mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}) &= \mathbb{E}(x_{\pi_1} - \mathbb{E}x_{\pi_1})\mathbb{E}(x_{\pi_2} - \mathbb{E}x_{\pi_2})\mathbb{E}y_{\pi_1}y_{\pi_2} \\ &\quad + \mathbb{E}x_{\pi_1}\mathbb{E}x_{\pi_2}\mathbb{E}(y_{\pi_1} - \mathbb{E}y_{\pi_1})\mathbb{E}(y_{\pi_2} - \mathbb{E}y_{\pi_2}) \\ &= 0. \end{aligned}$$

It thus follows that if  $(\pi_1, \pi_2) \notin K_{h,2}^{L_X} \cup K_{h,2}^{L_Y}$ , then

$$\mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2}) = 0.$$

Now, because of Assumption (A1),  $\mathbb{E}(z_{\pi_1} - \mathbb{E}z_{\pi_1})(z_{\pi_2} - \mathbb{E}z_{\pi_2})$  is bounded uniformly across all possible pairs of circuits and thus, from (3.4) and the bound (3.3), we conclude that

$$\text{Var}(\beta_h(n^{-1/2}Z_n)) = O\left(\frac{n^{h+1}}{n^{h+2}}\right) = O(n^{-1}).$$

This completes the proof.  $\square$

*Remark 3.1.* The verification of Condition (ii') of Lemma 2.1 requires more subtle combinatorial analysis. We have included it in Appendix B.

We now show how Carleman's condition can be checked easily under Property B provided that the word limit exists for every pair-matched word  $(w, w')$ . Let

$$\Pi_L^*(w) = \{\pi \mid w[i] = w[j] \Rightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j)), \text{ for all indices } (i, j)\}.$$

Clearly,  $\Pi_L(w) \subseteq \Pi_L^*(w)$ . Define

$$\Pi_Z^*(w, w') := \Pi_X^*(w) \cap \Pi_Y^*(w').$$

This also satisfies  $\Pi_Z(w, w') \subseteq \Pi_Z^*(w, w')$  and the set  $\Pi_Z^*(w, w') \setminus \Pi_Z(w, w')$  is contained in  $C_{h,3+}^{L_X} \cup C_{h,3+}^{L_Y}$  which means by Lemma 3.1 that

$$\lim_n n^{-(1+\frac{h}{2})} \#(\Pi_Z^*(w, w') \setminus \Pi_Z(w, w')) = 0, \text{ for each } (w, w') \in \mathcal{W}_{2k}^2.$$

Therefore, if  $\lim_n n^{-(1+\frac{h}{2})} \# \Pi_Z(w, w')$  exists, we have

$$\lim_n n^{-(1+\frac{h}{2})} \# \Pi_Z(w, w') = \lim_n n^{-(1+\frac{h}{2})} \# \Pi_Z^*(w, w') = p_Z(w, w'), \text{ say.}$$

**Theorem 3.1.** *Suppose that  $L_X$  and  $L_Y$  satisfy Property B and the input sequences satisfy Assumption (A1). If the limit  $p_Z(w, w')$  exists for every pair of pair-matched words  $(w, w')$  then Condition (i) of Lemma 2.1 holds and the limit moments satisfy Condition (iii) of Lemma 2.1 and hence are the moments of the LSD. The limit law in that case is sub-Gaussian. If the input sequences satisfy Assumptions (A2) or (A3) and the link functions satisfy Conditions 2.1, then the same LSD continues to hold.*

*Proof.* By the developments so far,

$$\begin{aligned} \lim_n \mathbb{E}(\beta_{2k}(F^{n^{-1/2}} Z_n)) &= \sum_{(w, w') \in \mathcal{W}_{2k}^2} \lim_n n^{-(1+\frac{h}{2})} \# \Pi_Z(w, w') \\ &= \sum_{(w, w') \in \mathcal{W}_{2k}^2} p_Z(w, w'). \end{aligned}$$

Then the moments of the LSD would be given by

$$\beta_h = \begin{cases} 0, & \text{if } h \text{ is odd} \\ \sum_{(w, w') \in \mathcal{W}_{2k}^2} p_Z(w, w'), & \text{if } h = 2k. \end{cases} \quad (3.7)$$

Now note that for any  $w \in \mathcal{A}_{2k}$ ,

$$\Pi_X(w) = \bigcup_{w' \in \mathcal{A}_{2k}} \Pi_Z(w, w').$$

Therefore, for  $w \in \mathcal{W}_{2k}$  we have

$$\bigcup_{w' \in \mathcal{W}_{2k}} \Pi_Z(w, w') \subseteq \Pi_X(w),$$

which implies that

$$\# \left( \bigcup_{w' \in \mathcal{W}_{2k}} \Pi_Z(w, w') \right) \leq \# \Pi_X(w).$$

This means that for  $(w, w') \in \mathcal{W}_{2k}^2$  we have

$$\sum_{w' \in \mathcal{W}_{2k}} p_Z(w, w') \leq p_X(w),$$

and therefore

$$\sum_{(w, w') \in \mathcal{W}_{2k}^2} p_Z(w, w') \leq \sum_{w \in \mathcal{W}_{2k}} p_X(w).$$

Since  $X$  and  $Y$  play a symmetric role in  $Z$ , we have furthermore

$$\sum_{(w, w') \in \mathcal{W}_{2k}^2} p_Z(w, w') \leq \min \left\{ \sum_{w \in \mathcal{W}_{2k}} p_X(w), \sum_{w' \in \mathcal{W}_{2k}} p_Y(w') \right\}. \quad (3.8)$$

Recall that the matrices are assumed to satisfy Property B. Also note that the number of pair-matched words of length  $2k$  equals  $\frac{(2k)!}{2^k k!}$ . Then it is easy to see that

$$\beta_{2k} \leq \frac{(2k)!}{2^k k!} \Delta^k, \quad (3.9)$$

where  $\Delta := \min\{\Delta_{L_X}, \Delta_{L_Y}\}$ . This guarantees that  $\{\beta_h\}$  satisfies Carleman's condition and the limit law is sub-Gaussian.

Suppose now that the input sequences satisfy Assumption (A2) and the link functions satisfy Conditions 2.1. Then, by appropriate truncation of the input variables and strong law of large numbers, one can reduce that case to the case where Assumption (A1) holds. We omit the tedious details which are similar to the case for a single matrix (see, for example, [Bose and Sen \[2008\]](#)).

If the input sequences satisfy Assumption (A3), then all moments are bounded and all the moment calculations and bounds used so far go through. Again, we omit the details. Finally, note that by Remark 3.1 the LSD exists in the almost sure sense. This completes the proof.  $\square$

## 4 Some general results

Note that by Theorem 3.1, the LSD will exist if the limit  $p(w, w')$  exists for each pair-matched word pair  $(w, w')$ . In this section we shall consider several types of  $\{X_n\}$ ,  $\{Y_n\}$  and establish general results on the LSD of  $n^{-1/2}Z_n$ . We assume that all input sequences satisfy Assumption (A1), (A2) or (A3). But as discussed, we can work under Assumption (A1). We first establish an invariance theorem.

**Theorem 4.1.** *Suppose that  $L_X$  satisfies Property B and Conditions (2.1). Suppose  $p_X(w)$  exists for each word so that  $\mathcal{L}_X$  exists. Also suppose that there is a transformation  $\rho$  such that  $L_Y = \rho \circ L_X$  and  $L_Y$  also satisfies the above conditions. Then the LSD of  $\{n^{-1/2}Z_n\}$  exists almost surely and equals  $\mathcal{L}_X$ .*

*Proof.* It is enough to show that  $\{n^{-1/2}Z_n\}$  has the same limiting moment sequence as  $\{n^{-1/2}X_n\}$  and by the theory developed earlier it suffices to look at the even moments only. Suppose that  $w \in \mathcal{W}_{2k}$  and  $\pi \in \Pi_X^*(w)$ . Then

$$\begin{aligned} w[i] = w[j] &\Rightarrow L_X(\pi(i-1), \pi(i)) = L_X(\pi(j-1), \pi(j)) \\ &\Rightarrow \rho \circ L_X(\pi(i-1), \pi(i)) = \rho \circ L_X(\pi(j-1), \pi(j)), \\ \text{i.e., } L_Y(\pi(i-1), \pi(i)) &= L_Y(\pi(j-1), \pi(j)). \end{aligned}$$

Therefore  $\pi \in \Pi_Y^*(w)$  which means that  $\Pi_X^*(w) \subseteq \Pi_Y^*(w)$ . Therefore

$$\Pi_X^*(w) \cap \Pi_Y^*(w) = \Pi_X^*(w).$$

This implies that  $p_Z(w, w) = p_X(w)$ . But then

$$p_X(w) = p_Z(w, w) \leq \sum_{w' \in \mathcal{W}_{2k}} p_Z(w, w') \leq p_X(w),$$

i.e., for each  $w \in \mathcal{W}_{2k}$ ,

$$\sum_{w' \in \mathcal{W}_{2k}} p_Z(w, w') = p_X(w).$$

Therefore, one has

$$\begin{aligned} \beta_{2k}^Z &= \lim_n \mathbb{E}(\beta_{2k}(F^{n^{-1/2}}Z_n)) = \sum_{(w, w') \in \mathcal{W}_{2k} \times \mathcal{W}_{2k}} p_Z(w, w') \\ &= \sum_{w \in \mathcal{W}_{2k}} \sum_{w' \in \mathcal{W}_{2k}} p_Z(w, w') \\ &= \sum_{w \in \mathcal{W}_{2k}} p_X(w) \\ &= \beta_{2k}^X. \end{aligned}$$

This completes the proof. □

*Remark 4.1.* It is clear from the proof of Theorem 4.1 that we may let  $\rho$  depend on  $n$ . Indeed, all our arguments are for a fixed  $n$ . For the sake of brevity, we shall continue using  $\rho$  instead of  $\rho_n$ .

**Example 4.1.** *Suppose  $X_n = W_n$  and  $L_Y$  satisfies Property B and Conditions (2.1). Then the LSD of  $\{n^{-1/2}W_n \odot Y_n\}$  is the semi-circular law  $\mathcal{L}_W$  almost surely. In particular,  $Y_n$  could be any one among  $T_n$ ,  $H_n$ ,  $RC_n$ ,  $SC_n$  or  $DH_n$ .*

One interesting case is when  $X_n$  is a  $\pm 1$  Bernoulli (with  $p = 0.5$ ) Wigner matrix. In that case the Schur-Hadamard product may be interpreted as a randomly censored patterned matrix. Thus, for example, a randomly ( $\pm 1$ ) censored Toeplitz matrix will have the semi-circular law as its LSD. This is the result of [Beckwith et al. \[2011\]](#) in the  $p = 0.5$  case.

**Example 4.2.** Suppose that  $X_n = T_n$ . Also suppose that

$$L_Y(i, j) = \rho(|i - j|), \text{ for all } (i, j)$$

for some function  $\rho$  and satisfies Property B and Conditions (2.1). Then the LSD of  $\{n^{-1/2}T_n \odot Y_n\}$  is  $\mathcal{L}_T$ .

In particular, if  $Y_n = SC_n$ , then the LSD of  $\{n^{-1/2}T_n \odot SC_n\}$  is  $\mathcal{L}_T$ .

**Example 4.3.** Suppose that  $X_n = H_n$ . Also suppose that

$$L_Y(i, j) = \rho(i + j), \text{ for all } (i, j)$$

for some function  $\rho$  and satisfies Property B and Conditions (2.1). Then the LSD of  $\{n^{-1/2}H_n \odot Y_n\}$  is  $\mathcal{L}_H$ .

The link functions of the Reverse Circulant and the Doubly Symmetric Hankel matrices have the form  $\rho(i + j)$ . Therefore, if we take  $Y_n = RC_n$  or  $DH_n$ , we can conclude that the LSD of  $\{n^{-1/2}H_n \odot Y_n\}$  is  $\mathcal{L}_H$ .

Similarly, since  $L_{DH}$  is of the form  $\rho(L_{RC})$ , we conclude that the LSD of  $\{n^{-1/2}RC_n \odot DH_n\}$  is  $\mathcal{L}_{RC}$ .

*Remark 4.2.* Note that Theorem 4.1 gives us all the rows of Table 2 except the second one.

The next natural question is what happens when we transform both the link functions  $L_X$  and  $L_Y$ . We first consider the case where we have a single sequence  $X_n$  and its link function is transformed.

**Proposition 4.1.** Suppose the link function  $L_X$  satisfies Property B and Conditions (2.1). Suppose also that  $p_X(w)$  exists for each word so that  $\mathcal{L}_X$  exists. Suppose  $L_Y := \rho \circ L_X$  where  $\rho$  is an injective transformation. Then the LSD of  $\{n^{-1/2}Y_n\}$  exists and equals  $\mathcal{L}_X$ .

*Proof.* We first observe that  $L_Y$  also satisfies Property B. To prove this, for  $1 \leq k \leq n$  and  $t \in \text{range}(L_X)$ , define  $D_{k,t}^X = \{l \mid 1 \leq l \leq n, L_X(k, l) = t\}$  and similarly define  $D_{k,t}^Y$ . Note that

$$L_X(k, l) = t \Leftrightarrow L_Y(k, l) = \rho(t).$$

Thus  $D_{k,t}^X = D_{k,\rho(t)}^Y$ . Since  $\rho$  is injective and by definition  $\text{range}(L_Y) = \rho(\text{range}(L_X))$ , we have

$$\begin{aligned} \Delta_{L_X} &= \sup_n \sup_{t \in \text{range}(L_X)} \sup_{1 \leq k \leq n} \#D_{k,t}^X \\ &= \sup_n \sup_{\rho(t) \in \text{range}(L_Y)} \sup_{1 \leq k \leq n} \#D_{k,\rho(t)}^Y \\ &= \Delta_{L_Y}. \end{aligned}$$

It now suffices to show that  $\Pi_X^*(w) = \Pi_Y^*(w)$  for each  $w \in \mathcal{W}_{2k}$ . From the proof of Theorem 4.1, it follows that  $\Pi_X^*(w) \subseteq \Pi_Y^*(w)$ . To show the other way, suppose that  $\pi \in \Pi_Y^*(w)$ . Then

$$\begin{aligned} w[i] &= w[j] \\ \Rightarrow L_Y(\pi(i-1), \pi(i)) &= L_Y(\pi(j-1), \pi(j)) \\ \Rightarrow \rho \circ L_X(\pi(i-1), \pi(i)) &= \rho \circ L_X(\pi(j-1), \pi(j)) \\ \Rightarrow L_X(\pi(i-1), \pi(i)) &= L_X(\pi(j-1), \pi(j)) \quad (\text{by injectivity of } \rho). \end{aligned}$$

Thus  $\pi \in \Pi_X^*(w)$ . Finally, we note that because of the injectivity of  $\rho$ , we have  $k_n^X = k_n^Y$  and  $\alpha_n^X = \alpha_n^Y$ , so that  $L_Y$  satisfies Conditions (2.1). The proof is now complete.  $\square$

**Example 4.4.** Take  $\rho(i, j) = a^{i \wedge j} b^{i \vee j}$ , where  $a$  and  $b$  are coprime positive integers, which is injective and compose it with the Wigner link function to obtain the link function  $L(i, j) = a^{i \wedge j} b^{i \vee j}$ . Then the LSD of the corresponding patterned random matrix is the semi-circular law. Similarly, the patterned random matrix with the link function  $L(i, j) = (i - j)^2$  has the same LSD as the Toeplitz matrix.

The following proposition shows that if both  $L_X$  and  $L_Y$  are transformed via injective maps, then the LSD of their Schur-Hadamard product is preserved.

**Proposition 4.2.** Suppose  $X_n$  and  $Y_n$  are independent patterned matrices where the link functions  $L_X$  and  $L_Y$  satisfy Property B and Conditions 2.1. Suppose  $p_Z(w, w')$  exists for each pair-matched word-pair  $(w, w')$  so that the LSD of  $\{n^{-1/2} X_n \odot Y_n\}$  exists. Suppose  $\rho_1$  and  $\rho_2$  are injective transformations and  $L_U = \rho_1 \circ L_X$  and  $L_V = \rho_2 \circ L_Y$ . If  $U_n$  and  $V_n$  are independent patterned matrices with link functions  $L_U$  and  $L_V$  respectively, then the LSD of  $\{n^{-1/2} U_n \odot V_n\}$  exists and is same as that of  $\{n^{-1/2} X_n \odot Y_n\}$ .

*Proof.* From the proof of Proposition 4.1 we see that both  $L_U$  and  $L_V$  satisfy Property B and Conditions 2.1 and we have  $\Pi_X^*(w) = \Pi_U^*(w)$  and  $\Pi_Y^*(w) = \Pi_V^*(w)$  for each pair matched word  $w$ . Then, for each word-pair  $(w, w')$ , one has  $\Pi_X^*(w) \cap \Pi_Y^*(w') = \Pi_U^*(w) \cap \Pi_V^*(w')$ . This completes the proof.  $\square$

**Example 4.5.** In view of Examples 4.1 and 4.4 we conclude from Proposition 4.2 that the LSD of the Schur-Hadamard product sequence  $\{n^{-1/2} U_n \odot V_n\}$  where  $L_U(i, j) = a^{i \wedge j} b^{i \vee j}$  and  $L_V(i, j) = (i - j)^2$  is the semi-circular law.

*Remark 4.3.* If we drop the assumption of injectivity, all we can say is that  $\Pi_X^*(w) \cap \Pi_Y^*(w') \subseteq \Pi_U^*(w) \cap \Pi_V^*(w')$ . So, if the LSD of  $\{n^{-1/2} U_n \odot V_n\}$  exists, its moments will dominate the moments of the LSD of  $\{n^{-1/2} X_n \odot Y_n\}$ .

## 5 Toeplitz and Hankel

Theorem 4.1 does not cover the situation where  $L_Y$  is not a function of  $L_X$ , e.g., the case where  $X_n$  is Toeplitz and  $Y_n$  is Hankel (see the second row of Table 2). In order to proceed further we need the concept of *Catalan* words from Bose and Sen [2008].

A Catalan word of length 2 is just a double letter  $aa$ . In general, a Catalan word of length  $2k$ ,  $k > 1$ , is a word  $w \in \mathcal{W}_{2k}$  containing a double letter such that if one deletes the double letter the reduced word becomes a Catalan word of length  $2k-2$ . For example,  $abba$ ,  $aabbcc$ ,  $abccbdda$  are Catalan words whereas  $abab$ ,  $abccab$ ,  $abcddcab$  are not. The set of all Catalan word of length  $2k$  will be denoted by  $\mathcal{C}_{2k}$ . There is a bijection between Catalan words and non-crossing pair partitions of the set  $\{1, 2, \dots, 2k\}$  whence it follows that

$$\#\mathcal{C}_{2k} = \frac{1}{k+1} \binom{2k}{k}, \quad (5.1)$$

the ubiquitous Catalan number from combinatorics.

By the theory developed in Section 3, it suffices to compute  $p_Z(w, w')$  for different combination of word pairs  $(w, w') \in \mathcal{W}_{2k}^2$ . Note that  $\pi \in \Pi_X^*(w) \cap \Pi_Y^*(w')$  means that we have exactly  $2k$  constraints on the vertices  $\pi(i)$ , with each word giving rise to  $k$  constraints. To elaborate, each  $w$ -match  $(i, j)$  gives rise to the restriction

$$L_X(\pi(i-1), \pi(i)) = L_X(\pi(j-1), \pi(j)),$$

and each  $w'$ -match  $(k, l)$  gives rise to the restriction

$$L_Y(\pi(k-1), \pi(k)) = L_Y(\pi(l-1), \pi(l)).$$

Thus we expect that if  $w \neq w'$  and the two links functions  $L_X$  and  $L_Y$  behave nicely, then we will have more than  $k$  independent constraints so that  $\#\Pi_X^*(w) \cap \Pi_Y^*(w') = O(n^k)$  and a fortiori  $p_Z(w, w') = 0$ . We shall call two link functions  $L_X$  and  $L_Y$ , which satisfy Property B, *compatible* if, for  $w \neq w'$ , we have  $p_Z(w, w') = 0$ . We shall also write  $(L_X, L_Y) \rightsquigarrow L_W$  if

$$p_Z(w, w) = \begin{cases} 1, & \text{if } w \in \mathcal{C}_{2k} \\ 0, & \text{otherwise.} \end{cases}$$

**Proposition 5.1.** *Suppose  $L_X$  and  $L_Y$  are compatible and  $(L_X, L_Y) \rightsquigarrow L_W$ . Then the LSD of  $\{n^{-1/2}X_n \odot Y_n\}$  is the semi-circular law.*

*Proof.* We have

$$\begin{aligned} \beta_{2k}^Z &= \sum_{(w, w') \in \mathcal{W}_{2k}^2} p_Z(w, w') \\ &= \sum_{w \in \mathcal{C}_{2k}} p_Z(w, w) = \#\mathcal{C}_{2k}. \end{aligned}$$



This completes the proof because the  $2k$ -th moment of the semi-circular law is  $\#\mathcal{C}_{2k}$ .  $\square$

We write  $(L_X, L_Y) \Rightarrow L_W$  if  $L_X$  and  $L_Y$  together determine the Wigner link function  $L_W$  in the sense that  $L_X(i, j) = L_X(k, l)$  and  $L_Y(i, j) = L_Y(k, l)$  together imply that  $L_W(i, j) = L_W(k, l)$ .

**Lemma 5.1.** *If  $(L_X, L_Y) \Rightarrow L_W$ , then  $(L_X, L_Y) \rightsquigarrow L_W$ .*

*Proof.* Suppose that  $\pi \in \Pi_X^*(w) \cap \Pi_Y^*(w)$ . Then

$$\begin{aligned} w[i] &= w[j] \\ \Rightarrow L_X(\pi(i-1), \pi(i)) &= L_X(\pi(j-1), \pi(j)) \text{ and } L_Y(\pi(i-1), \pi(i)) = L_Y(\pi(j-1), \pi(j)) \\ \Rightarrow L_W(\pi(i-1), \pi(i)) &= L_W(\pi(j-1), \pi(j)). \end{aligned}$$

Therefore  $\pi \in \Pi_W^*(w)$  and as a consequence  $\Pi_X^*(w) \cap \Pi_Y^*(w) \subseteq \Pi_W^*(w)$ . The other inclusion is always true because  $L_X$  and  $L_Y$  are symmetric link functions so that  $\Pi_W^*(w) \subseteq \Pi_X^*(w)$  and  $\Pi_W^*(w) \subseteq \Pi_Y^*(w)$ . Therefore  $\Pi_X^*(w) \cap \Pi_Y^*(w) = \Pi_W^*(w)$  which means that we have

$$p_Z(w, w) = p_W(w),$$

for each  $w \in \mathcal{W}_{2k}$ . As for the Wigner matrix

$$p_W(w) = \begin{cases} 1, & \text{if } w \in \mathcal{C}_{2k} \\ 0, & \text{otherwise,} \end{cases}$$

the proof is now complete.  $\square$

We shall establish below that the Schur-Hadamard product of Toeplitz and Hankel has the semi-circular LSD by verifying the conditions of Proposition 5.1. See Figure 1.

We first make a simplification. Let  $s(i) = \pi(i) - \pi(i-1)$ . Define

$$\Pi'(w) = \{\pi \mid w[i] = w[j] \Rightarrow s(i) + s(j) = 0\}.$$

From Bose and Sen [2008] it is known that for the Toeplitz matrix,

$$p(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi'(w).$$

Therefore, we need only look at  $\Pi'_X(w) \cap \Pi_Y^*(w')$  and if the limit exists we have

$$p_Z(w, w') = \lim_n \frac{1}{n^{1+k}} \#\Pi'_X(w) \cap \Pi_Y^*(w').$$

We shall use this in the following lemma.

**Lemma 5.2.**  *$L_T$  and  $L_H$  are compatible.*

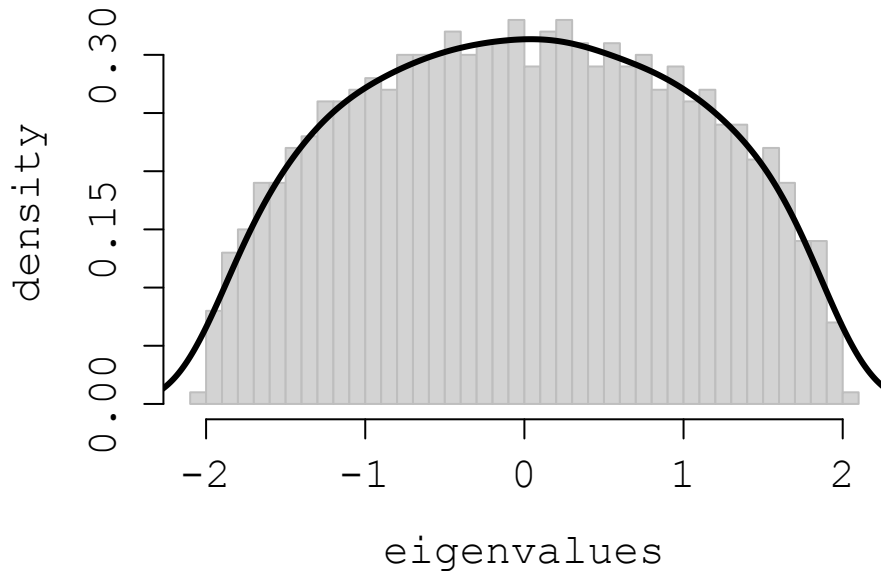


Figure 1: Histogram and kernel density estimate for the ESD of  $n^{-1/2}X_n \odot Y_n$ , where  $n = 1000$ ,  $X_n$  is a random Toeplitz matrix,  $Y_n$  is a random Hankel matrix, they are independent and have  $\mathcal{N}(0, 1)$  entries.

*Proof.* We shall show that if  $w \neq w'$ , then  $\#\Pi'_X(w) \cap \Pi_Y^*(w') = O(n^k)$ . This would imply that

$$p_Z(w, w') = 0, \text{ if } w \neq w'.$$

To do this it is enough to show that in addition to the  $k$  constraints on the choices of  $\pi$  arising from the Toeplitz link function (or the Hankel link function), there is at least one more additional constraint. Note that for Toeplitz or Hankel link functions, the natural constraints arising from matches enable one to express each non-generating vertex  $\pi(j)$  as a linear combination of the generating vertices  $\pi(i)$  preceding it (i.e.,  $i < j$ ). We shall show that if we combine the  $2k$  constraints corresponding to  $(w, w')$ , then we can write some generating vertex  $\pi(i)$  as a linear combination of the preceding generating vertices. This will be the extra constraint we are seeking.

Consider the positions where new letters (letters appearing for the first time) appear. Note that the positions of the new letters fix their pattern and also fix the position of the old letters but not their pattern. Suppose the positions of the new letters are not all same in  $w$  and  $w'$ . Let  $i$  be the *first* place where a new letter appears in (say)  $w$  but an old letter appears in  $w'$ . So  $\pi(i)$  is a generating vertex for  $\Pi_T(w)$ , but for  $\Pi_H(w')$  it is non-generating, so that we can express  $\pi(i)$  as a linear

combination of generating vertices  $\pi(k)$ ,  $k < i$ ,  $k \in S_H$ . but note that prior to  $i$  the generating vertices in  $S_T$  and  $S_H$  are same (indeed  $i$  is the first position where there is a difference). Therefore, we can express  $\pi(i)$  as a linear combination of generating vertices  $\pi(k)$ ,  $k < i$ ,  $k \in S_T$ , which is an extra constraint.

Now suppose that all the new letters appear at the same positions in  $w$  and  $w'$  so that  $S_T = S_H = S$ , say. Now, since  $w \neq w'$ , there exists  $j$  such that there are  $i, i^* \in S$  both less than  $j$  with  $i \sim^T j$  and  $i^* \sim^H j$ . Assume that  $j$  is the first letter of this type. Further without loss of generality we may assume that  $i^* < i$ . Then we have the following two constraints:

$$\pi(i) - \pi(i-1) = \pi(j-1) - \pi(j), \text{ and} \quad (5.2)$$

$$\pi(i^*) + \pi(i^*-1) = \pi(j) + \pi(j-1). \quad (5.3)$$

Eliminating  $\pi(j)$  from these two constraints we arrive at

$$2\pi(j-1) = \pi(i) - \pi(i-1) + \pi(i^*) + \pi(i^*-1). \quad (5.4)$$

**Case I:** ( $i = j-1$ ). In this case (5.4) becomes

$$\pi(i) = -\pi(i-1) + \pi(i^*) + \pi(i^*-1),$$

which is an additional constraint because we are being able to express the generating vertex  $\pi(i)$  as a linear combination of generating vertices  $\pi(k)$  with  $k < i$ .

**Case II:** ( $i < j-1$ ). If  $\pi(j-1)$  is a generating vertex, then again we have an extra constraint because via (5.4) we are able to express the generating vertex  $\pi(j-1)$  as a linear combination of generating vertices  $\pi(k)$  with  $k < j-1$ . So suppose that  $\pi(j-1)$  is non-generating. Then there exists  $i_1 \in S$  such that  $i_1 \sim^T (j-1)$  and  $i_1 \sim^H (j-1)$  (recall that  $j$  is assumed to be the first index where the words  $w$  and  $w'$  differ). This implies that we have

$$\pi(i_1) - \pi(i_1-1) = \pi(j-2) - \pi(j-1), \text{ and}$$

$$\pi(i_1) + \pi(i_1-1) = \pi(j-2) + \pi(j-1),$$

which simplify to

$$\pi(i_1) = \pi(j-2), \text{ and}$$

$$\pi(i_1-1) = \pi(j-1).$$

Thus (5.4) becomes

$$2\pi(i_1-1) = \pi(i) - \pi(i-1) + \pi(i^*) + \pi(i^*-1). \quad (5.5)$$

If  $i_1-1 \leq i$ , using (5.5) we can express  $\pi(i)$  as a linear combination of generating vertices  $\pi(k)$  with  $k < i$  thus giving rise to an extra constraint. On the other hand

if  $i_1 - 1 > i$  and  $i_1 - 1 \in S$  then (5.5) gives an extra constraint where  $\pi(i_1 - 1)$  is expressed as a linear combination of generating vertices  $\pi(k)$  with  $k < i_1 - 1$ . Finally, if  $i_1 - 1 > i$  and  $i_1 - 1 \notin S$ , then, by the same argument as above with the role of  $j - 1$  being played by  $i_1 - 1$ , we can find  $i_2 \in S$  such that  $i_2 \sim^T (i_1 - 1)$  and  $i_2 \sim^H (i_2 - 1)$  and so on. It is clear that if we continue this procedure, then at some point we will obtain  $i_m \in S$ ,  $m \geq 1$  such that  $i_m \sim^T (i_{m-1} - 1)$  and  $i_m \sim^H (i_{m-1} - 1)$  and  $i_m - 1 \leq i$  and so we will be able to express  $\pi(i)$  as a linear combination of generating vertices  $\pi(k)$  with  $k < i$ , thus obtaining an extra constraint.  $\square$

**Lemma 5.3.**  $(L_T, L_H) \rightsquigarrow L_W$ .

*Proof.* As  $(L_T, L_H) \Rightarrow L_W$ , Lemma 5.1 directly applies. However, we give here an alternate argument that applies to some cases where  $(L_X, L_Y) \not\Rightarrow L_W$  (for example, note that  $(L_T, L_{RC}) \not\Rightarrow L_W$  but one can show that  $(L_T, L_{RC}) \rightsquigarrow L_W$ , see Remark 5.1 below).

Fix  $w \in \mathcal{C}_{2k}$ . Suppose that we have a double letter at position  $i$ , i.e.,  $w[i] = w[i + 1]$ . Suppose  $\pi \in \Pi_Y^*(w)$ . Then we have

$$\pi(i - 1) + \pi(i) = \pi(i) + \pi(i + 1).$$

So  $\pi(i - 1) = \pi(i + 1)$ . This implies that

$$s(i) = \pi(i) - \pi(i - 1) = \pi(i) - \pi(i + 1) = -s(i + 1).$$

Now, deleting this double letter, i.e., identifying  $\pi(i - 1)$  and  $\pi(i)$ , we are left with a Catalan word  $\hat{w}$  of length  $2k - 2$  and the reduced circuit  $\hat{\pi} \in \Pi_Y^*(\hat{w})$ . It has a double letter and we can repeatedly use the above argument until the whole word is emptied. What this argument gives is this:  $w[i] = w[j] \Rightarrow s(i) + s(j) = 0$ . So  $\pi \in \Pi_X'(w)$ . Therefore  $\Pi_Y^*(w) \subseteq \Pi_X'(w)$ . So  $\Pi_X'(w) \cap \Pi_Y^*(w) = \Pi_Y^*(w)$  and therefore  $p_Z(w, w)$  exists and equals  $p_Y(w) = 1$ .

Now fix  $w \in \mathcal{W}_{2k}$ . Consider the match  $(i, j)$  where  $j$  is the position of the first old letter. We have

$$\pi(i) - \pi(i - 1) = \pi(j - 1) - \pi(j), \text{ and}$$

$$\pi(i) + \pi(i - 1) = \pi(j - 1) + \pi(j),$$

which simplify to

$$\pi(i) = \pi(j - 1), \text{ and} \tag{5.6}$$

$$\pi(i - 1) = \pi(j). \tag{5.7}$$

But by definition of  $j$ ,  $\pi(j - 1)$  is generating, so (5.6) is a new constraint unless  $i = j - 1$ , in which case  $(i, j)$  is a double letter and (5.7) becomes  $\pi(i - 1) = \pi(i + 1)$ , which is an automatic constraint in both  $\Pi_X'(w)$  and  $\Pi_Y^*(w)$ . Delete this double letter and apply the above argument on the reduced word. Clearly, if  $w$  is non-Catalan, at some point we will be left with a non-empty word with no double letters, thus getting an extra constraint and therefore we will have  $p_Z(w, w) = 0$ .  $\square$

*Remark 5.1.* If we take  $X_n$  as Symmetric Circulant, then Lemmas 5.2 and 5.3 continue to hold. Indeed, if for a pair matched word  $w$  we define

$$\Pi'_X(w) = \{\pi \mid w[i] = w[j] \Rightarrow s(i) + s(j) = 0, \pm n\},$$

then, from Bose and Sen [2008], we know that

$$p_X(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi'_X(w) = 1.$$

One can readily see that the proof of Lemmas 5.2 and 5.3 goes through in this case with minor modifications. Similarly, we can take  $Y_n$  as Reverse Circulant or Doubly Symmetric Hankel. Therefore, the conclusion of Theorem 5.1 hold in these cases as well and thus we completely obtain the second row of Table 2.

## Appendix A Two counting lemmas

Recall that  $K_{h,t}^L$  is the set of  $t$ -tuples of circuits  $(\pi_1, \dots, \pi_t)$  of length  $h$  such that they are jointly and across  $L$ -matched. If  $L$  satisfies Property  $B$ , then Lemma 2(a) of Bose and Sen [2008] says that

$$\#K_{h,4}^L = O(n^{2h+2}). \quad (\text{A.1})$$

The arguments of Bose and Sen [2008] are adaptations of those of Bryc et al. [2006] who proved this estimate for Toeplitz and Hankel matrices. One can modify these arguments to accommodate other values of  $t$ . For the reader's convenience we provide a proof for  $t = 2$  which we have used in the proof of Lemma 3.2 and will be using in the proof of Lemma B.1. We also state the version for  $t = 3$  without proof as it will be needed while proving Lemma B.1.

**Lemma A.1.** *If  $L$  satisfies Property  $B$ , then*

$$\begin{aligned} \#K_{h,2}^L &= O(n^{h+1}), \text{ and} \\ \#K_{h,3}^L &= O(n^{\lfloor \frac{3h}{2} \rfloor + 2}). \end{aligned}$$

*Proof.* Consider all circuits  $(\pi_1, \pi_2)$  of length  $h$  which are jointly  $L$ -matched and across  $L$ -matched. Consider all possible edges  $(\pi_j(i-1), \pi_j(i))$ ,  $1 \leq j \leq 2$  and  $1 \leq i \leq h$ . Since the circuits are jointly and across  $L$ -matched, there are at most  $h$  distinct  $L$ -values in these  $2h$  edges.

Note that the number of partitions of the  $2h$  edges into distinct groups of  $L$ -matching edges, with at least two edges in each group, is independent of  $n$ . So, for a fixed integer  $1 \leq u \leq h$ , it is enough to establish the required estimate for the number of pairs of circuits for which there are exactly  $u$  distinct  $L$ -values.

First assume that  $1 \leq u \leq h-1$ . We count the total number of choices in the following way:

1. The generating vertices  $\pi_1(0), \pi_2(0)$  may be chosen in total  $n^2$  many ways.
2. Now arrange the values  $L(\pi_j(i-1), \pi_j(i))$ ,  $1 \leq j \leq 2$ ,  $1 \leq i \leq h$  from left to right, starting with  $\pi_1$  followed by  $\pi_2$ . Then the generating vertices  $\pi_j(i)$ , for which  $L(\pi_j(i-1), \pi_j(i))$  is the first one of the distinct  $L$ -values in this sequence, have at most  $n^u$  choices.
3. Having chosen these vertices, using Property B and  $L$ -matchings, the rest of the vertices in all the circuits may be chosen from left to right in at most  $(\Delta_L)^{2h-u-2}$  ways.

Now, since  $u \leq h-1$ , the total number of choices is bounded by

$$n^2 n^u (\Delta_L)^{2h-u-2} = O(n^{u+2}) = O(n^{h+1}).$$

Now consider the case  $u = h$ . Then each  $L$ -value is shared by exactly two edges. Now we seek to identify one generating vertex that has only finitely many choices.

By reordering the two circuits if necessary, we have an  $L$ -value that is assigned, as the first and only one, to exactly one edge, say  $(\pi_1(i-1), \pi_1(i))$  of  $\pi_1$ . Pick this  $L$ -value. The rest of the  $(u-1)$  generating vertices may be chosen in at most  $n^{u-1} = n^{h-1}$  ways. By the following dynamic construction of  $\pi_1$  we show that  $\pi_1(i)$  can have only finitely many choices:

Start with  $\pi_1(0)$  and choose  $\pi_1(j)$  till  $j \leq i-1$ , honouring the  $L$ -matches. Now start from the tail end of  $\pi_1$ , i.e., from  $\pi_1(h) = \pi_1(0)$  and choose the vertices  $\pi_1(j)$  in a right-to-left manner. When  $\pi_1(i+1)$  is chosen, since the  $L$ -value  $L(\pi_1(i), \pi_1(i+1))$  appears elsewhere, note that  $\pi_1(i)$  can have only finitely many choices. Thus the total number of choices is bounded by

$$n^2 n^{u-1} (\Delta_L)^{2h-u-2+1} = O(n^{u+1}) = O(n^{h+1}).$$

This completes the proof in the case  $t = 2$ . The proof in the case  $t = 3$  is an easy modification of the argument above and hence omitted.  $\square$

The following lemma will be repeatedly used in the verification of Condition (ii') of Lemma 2.1 in Lemma B.1 of Appendix B.

**Lemma A.2.** *Consider two  $h$ -circuits  $\pi_1$  and  $\pi_2$ . Suppose  $\pi_1$  is pair-matched with respect to  $L_X$  (which necessitates that  $h$  be even) and shares no  $L_X$  values with  $\pi_2$ . Also suppose that  $\pi_1$  and  $\pi_2$  share an  $L_Y$  value. Then, contingent on the event that  $\pi_2$  has been already chosen, one can choose  $\pi_1$  in  $O(n^{h/2})$  ways, honouring the stated constraints.*

*Proof.* Consider  $\pi_1$  from left to right. There is one and hence a first index  $i$  such that the  $L_Y$ -value  $L_Y(\pi_1(i-1), \pi_1(i))$  appears in  $\pi_2$ . Now consider the  $L_X$ -matches on  $\pi_1$ . Since  $\pi_1$  is pair-matched, it has  $h/2 + 1$  generating vertices and therefore in absence of any further constraints one can choose  $\pi_1$  in  $O(n^{h/2+1})$  ways. We shall

show that under the setup of the lemma one among these generating vertices has only finitely many choices. Note that we may assume without loss of generality that  $\pi_1(i)$  is a generating vertex with respect to  $L_X$  (indeed, otherwise we may start filling the circuit from right to left and define generating vertices according to that order to ensure that  $\pi_1(i)$  is generating). But now, since the value  $L_Y(\pi_1(i-1), \pi_1(i))$  is fixed, after choosing  $\pi_1(0), \dots, \pi_1(i-1)$  with respect to  $L_X$ , there are only finitely many choices left for (the generating vertex)  $\pi_1(i)$ . This completes the proof.  $\square$

## Appendix B Almost sure weak convergence

The proof of almost sure weak convergence is presented in the following lemma.

**Lemma B.1.** *Suppose  $L_X$  and  $L_Y$  satisfy Property B and the input sequences satisfy Assumption (A1). Then  $\{\beta_h(n^{-1/2}Z_n)\}$  satisfies Condition (ii') of Lemma 2.1 for any  $h$ .*

*Proof.* Using (A.1) in our context we have

$$\begin{aligned} \#(K_{h,4}^{L_X} \cup K_{h,4}^{L_Y}) &\leq \#K_{h,4}^{L_X} + \#K_{h,4}^{L_Y} \\ &= O(n^{2h+2}) + O(n^{2h+2}) \\ &= O(n^{2h+2}). \end{aligned} \tag{B.1}$$

Now write

$$\begin{aligned} \mathbb{E}[\beta_h(n^{-1/2}Z_n) - \mathbb{E}(\beta_h(n^{-1/2}Z_n))]^4 &= \mathbb{E}[n^{-1} \text{tr}(n^{-1/2}Z_n)^h - \mathbb{E}(n^{-1} \text{tr}(n^{-1/2}Z_n)^h)]^4 \\ & \tag{B.2} \end{aligned}$$

$$\begin{aligned} &= \frac{1}{n^{2h+4}} \mathbb{E}[\text{tr} Z_n^h - \mathbb{E} \text{tr} Z_n^h]^4 \\ &= \frac{1}{n^{2h+4}} \sum_{(\pi_1, \pi_2, \pi_3, \pi_4)} \mathbb{E}(\prod_{j=1}^4 z_{\pi_j} - \mathbb{E} z_{\pi_j}). \end{aligned}$$

Therefore, using decomposition (3.5) we have

$$\prod_{j=1}^4 (z_{\pi_j} - \mathbb{E} z_{\pi_j}) = \prod_{j=1}^4 ((x_{\pi_j} - \mathbb{E} x_{\pi_j})(y_{\pi_j} - \mathbb{E} y_{\pi_j}) + (y_{\pi_j} - \mathbb{E} y_{\pi_j})\mathbb{E} x_{\pi_j} + (x_{\pi_j} - \mathbb{E} x_{\pi_j})\mathbb{E} y_{\pi_j}).$$

If  $(\pi_1, \pi_2, \pi_3, \pi_4)$  are not jointly  $L_X$ -matched, then one of the circuits, say  $\pi_k$ , has an  $L_X$ -value which does not occur anywhere else. Therefore  $\mathbb{E} x_{\pi_k} = 0$ . So

$$z_{\pi_k} - \mathbb{E} z_{\pi_k} = x_{\pi_k} y_{\pi_k},$$

and

$$\prod_{j=1}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j}) = x_{\pi_k} y_{\pi_k} \prod_{\substack{j=1 \\ j \neq k}}^4 (x_{\pi_j} y_{\pi_j} - \mathbb{E}x_{\pi_j} \mathbb{E}y_{\pi_j}). \quad (\text{B.3})$$

Because of the independence of  $X_n$  and  $Y_n$  and of the input sequences we can conclude from this representation that

$$\mathbb{E} \prod_{j=1}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j}) = 0,$$

since the input  $X$ -variable corresponding to the single  $L_X$ -value appears in the product (B.3) inside  $x_{\pi_k}$  and is independent of every other term in the product. Therefore, in order to have a non-zero contribution,  $(\pi_1, \pi_2, \pi_3, \pi_4)$  have to be jointly  $L_X$ -matched and by the same argument jointly  $L_Y$ -matched.

Now suppose that  $(\pi_1, \pi_2, \pi_3, \pi_4)$  are jointly  $L_X$  as well as  $L_Y$ -matched but neither across  $L_X$ -matched nor across  $L_Y$ -matched. Then there is a circuit, say  $\pi_k$ , which is only *self*  $L_X$ -matched, i.e., none of its  $L_X$ -values is shared with those of the other circuits. Similarly, there is a circuit  $\pi_l$  that is only self  $L_Y$ -matched. Now note that  $(x_{\pi_k} - \mathbb{E}x_{\pi_k})$  is independent of  $(x_{\pi_j} - \mathbb{E}x_{\pi_j})$  for  $j \neq k$  and similarly  $(y_{\pi_l} - \mathbb{E}y_{\pi_l})$  is independent of  $(y_{\pi_j} - \mathbb{E}y_{\pi_j})$  for  $j \neq l$ . If  $k = l$  (which is always the case in the setup of Theorem 4.1), using these facts along with the independence of  $X_n$  and  $Y_n$  and the decomposition (3.5) we can write

$$\begin{aligned} \mathbb{E} \prod_{j=1}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j}) &= \mathbb{E}(x_{\pi_k} - \mathbb{E}x_{\pi_k}) \mathbb{E}(y_{\pi_k}) \mathbb{E} \left( \prod_{\substack{j=1 \\ j \neq k}}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j}) \right) \\ &\quad + \mathbb{E}(y_{\pi_k} - \mathbb{E}y_{\pi_k}) \mathbb{E}(x_{\pi_k}) \mathbb{E} \left( \prod_{\substack{j=1 \\ j \neq k}}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j}) \right) \\ &= 0. \end{aligned}$$

If  $k \neq l$  then  $\mathbb{E} \prod_{j=1}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j})$  is not necessarily 0. However, since, by Assumption (A1),  $\mathbb{E} \prod_{j=1}^4 (z_{\pi_j} - \mathbb{E}z_{\pi_j})$  is bounded uniformly across all possible quadruples, it suffices to prove an  $O(n^{2h+3-\delta})$  estimate,  $\delta > 0$ , on the number of quadruples of circuits in the  $k \neq l$  case. We shall prove such estimates (and we will not try to be optimal) in each of the following three possible cases:

**Case I.**  $\pi_1$  is self  $L_X$ -matched,  $(\pi_2, \pi_3, \pi_4)$  are across  $L_X$ -matched and  $\pi_2$  is self  $L_Y$ -matched,  $(\pi_1, \pi_3, \pi_4)$  are across  $L_Y$ -matched. By Lemma A.1, if we just consider the  $L_X$ -matches, then  $(\pi_2, \pi_3, \pi_4)$  can be chosen together in at most  $O(n^{\lfloor 3h/2 \rfloor + 2})$  many ways. So, if  $\pi_1$  has at least one edge of order  $\geq 3$ , then, by (3.2), we can choose  $\pi_1$  in  $O(n^{\lfloor (h+1)/2 \rfloor})$  ways. Thus, in this case, the total number choices for the



quadruples  $(\pi_1, \pi_2, \pi_3, \pi_4)$  is

$$\underbrace{O(n^{\lfloor \frac{3h}{2} \rfloor + 2})}_{(\pi_2, \pi_3, \pi_4)} \underbrace{O(n^{\lfloor \frac{h+1}{2} \rfloor})}_{\pi_1} = O(n^{2h + \frac{5}{2}}).$$

So we may assume that  $\pi_1$  is pair-matched with respect to  $L_X$  and by the same token  $\pi_2$  is pair-matched with respect to  $L_Y$ .

Choose  $(\pi_2, \pi_3, \pi_4)$  honouring the  $L_X$ -matches in  $O(n^{\lfloor 3h/2 \rfloor + 2})$  ways. Now  $\pi_1$  shares an  $L_Y$  value either with  $\pi_3$  or  $\pi_4$ , since  $(\pi_1, \pi_3, \pi_4)$  are across  $L_Y$ -matched. Therefore, by Lemma A.2 we can choose  $\pi_1$  in  $O(n^{h/2})$  ways. Therefore, the total number of choices for the quadruples  $(\pi_1, \pi_2, \pi_3, \pi_4)$  is

$$\underbrace{O(n^{\lfloor 3h/2 \rfloor + 2})}_{(\pi_2, \pi_3, \pi_4)} \underbrace{O(n^{\frac{h}{2}})}_{\pi_1} = O(n^{2h+2}).$$

**Case II.**  $\pi_1, \pi_2$  are self  $L_X$ -matched and  $(\pi_3, \pi_4)$  are across  $L_X$ -matched while  $\pi_3$  is self  $L_Y$ -matched and  $(\pi_1, \pi_2, \pi_4)$  are across  $L_Y$ -matched. Note that we may again assume that both  $\pi_1$  and  $\pi_2$  are pair-matched with respect to  $L_X$ , because otherwise upon choosing  $(\pi_3, \pi_4)$ , honouring the  $L_X$ -matches, in  $O(n^{h+1})$  ways (by Lemma A.1), we can choose both  $\pi_1$  and  $\pi_2$  in  $O(n^{\lfloor \frac{h+1}{2} \rfloor})$  ways with respect to  $L_X$ , so that the total number of choices becomes

$$\underbrace{O(n^{h+1})}_{(\pi_3, \pi_4)} \underbrace{O(n^{\lfloor \frac{h+1}{2} \rfloor})}_{\pi_2} \underbrace{O(n^{\lfloor \frac{h+1}{2} \rfloor})}_{\pi_1} = O(n^{2h+2}).$$

Choose  $(\pi_3, \pi_4)$ , honouring the  $L_X$ -matches, in  $O(n^{h+1})$  ways. Now, since  $(\pi_1, \pi_2, \pi_4)$  are across  $L_Y$ -matched, two possibilities might arise:

1.  $\pi_1$  and  $\pi_2$  both share an  $L_Y$ -value with  $\pi_4$ .
2.  $\pi_1, \pi_2$  share an  $L_Y$ -value and  $\pi_2, \pi_4$  share an  $L_Y$ -value.

In the first case, since we have already chosen  $\pi_4$ , by Lemma A.2  $\pi_1$  and  $\pi_2$  both can be chosen in  $O(n^{h/2})$  ways so that the total number of choices for the quadruples  $(\pi_1, \pi_2, \pi_3, \pi_4)$  is

$$\underbrace{O(n^{h+1})}_{(\pi_3, \pi_4)} \underbrace{O(n^{\frac{h}{2}})}_{\pi_2} \underbrace{O(n^{\frac{h}{2}})}_{\pi_1} = O(n^{2h+1}).$$

In the second case, again by Lemma A.2, we can choose  $\pi_2$  in  $O(n^{h/2})$  ways and thereafter  $\pi_1$  in  $O(n^{h/2})$  ways so that the total number of choices again becomes

$$\underbrace{O(n^{h+1})}_{(\pi_3, \pi_4)} \underbrace{O(n^{\frac{h}{2}})}_{\pi_2} \underbrace{O(n^{\frac{h}{2}})}_{\pi_1} = O(n^{2h+1}).$$

**Case III.**  $\pi_1, \pi_2$  are self  $L_X$ -matched and  $(\pi_3, \pi_4)$  are across  $L_X$ -matched while  $\pi_3, \pi_4$  are self  $L_Y$ -matched and  $(\pi_1, \pi_2)$  are across  $L_Y$ -matched. Once again we may

and will assume that both  $\pi_1$  and  $\pi_2$  are pair-matched with respect to  $L_X$ . Choose  $(\pi_3, \pi_4)$  honouring the  $L_X$  constraints in  $O(n^{h+1})$  ways. Now choose  $\pi_2$  in  $O(n^{h/2+1})$  ways honouring the  $L_X$  constraints. Since  $(\pi_1, \pi_2)$  are across  $L_Y$ -matched and  $\pi_2$  has been chosen, by Lemma A.2 we can choose  $\pi_1$  in  $O(n^{h/2})$  ways. Thus, in this case, the total number of choices for the quadruples  $(\pi_1, \pi_2, \pi_3, \pi_4)$  is

$$\underbrace{O(n^{h+1})}_{(\pi_3, \pi_4)} \underbrace{O(n^{\frac{h}{2}+1})}_{\pi_2} \underbrace{O(n^{\frac{h}{2}})}_{\pi_1} = O(n^{2h+2}).$$

All the other types of quadruples of circuits are contained in  $K_{h,4}^{L_X} \cup K_{h,4}^{L_Y}$ . Therefore, by what have been established so far, we conclude that

$$\mathbb{E}[\beta_h(n^{-1/2}Z_n) - \mathbb{E}(\beta_h(n^{-1/2}Z_n))]^4 = O(n^{-(1+\delta)}),$$

for some suitable  $\delta > 0$ , which completes the verification of Condition (ii') of Lemma 2.1.  $\square$

## 6 Acknowledgements

We thank the anonymous referees for their comments and for pointing us to important literature that we had missed.

## References

- Bai, Z. and L. Zhang (2007). Semicircle law for Hadamard products. *SIAM Journal on Matrix Analysis and Applications* 29(2), 473–495.
- Basu, R., A. Bose, S. Ganguly, and R. S. Hazra (2012). Joint convergence of several copies of different patterned random matrices. *Electron. J. Probab.* 17(82), 1–33.
- Basu, R., A. Bose, S. Ganguly, and R. Subhra Hazra (2012). Spectral properties of random triangular matrices. *Random Matrices. Theory and Applications* 1(3), 1250003, 22.
- Beckwith, O., V. Luo, S. J. Miller, K. Shen, and N. Triantafillou (2011). Distribution of eigenvalues of weighted, structured matrix ensembles. *arXiv preprint arXiv:1112.3719*.
- Bose, A., R. S. Hazra, and K. Saha (2010). Patterned random matrices and method of moments. In *Proceedings of the International Congress of Mathematicians, Hyderabad*, pp. 2203–2230.
- Bose, A., R. S. Hazra, and K. Saha (2011). Convergence of joint moments for independent random patterned matrices. *The Annals of Probability* 39(4), 1607–1620.

- Bose, A. and A. Sen (2008). Another look at the moment method for large dimensional random matrices. *Electron. J. Probab.* 13(21), 588–628.
- Bryc, W., A. Dembo, and T. Jiang (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *The Annals of Probability* 34(1), 1–38.
- Goldmakher, L., C. Khoury, S. J. Miller, and K. Ninsuwan (2013). On the spectral distribution of large weighted random regular graphs. *arXiv preprint arXiv:1306.6714*.